

Müller, Hans-Georg; Kepser, Matthias; Schallenger, Stefan  
**Getrennt- und Zusammenschreibung - Ein konsistentes Konstrukt? Erste Ergebnisse einer empirischen Studie**

*Didaktik Deutsch : Halbjahresschrift für die Didaktik der deutschen Sprache und Literatur* 23 (2018) 45, S. 74-93



Quellenangabe/ Reference:

Müller, Hans-Georg; Kepser, Matthias; Schallenger, Stefan: Getrennt- und Zusammenschreibung - Ein konsistentes Konstrukt? Erste Ergebnisse einer empirischen Studie - In: *Didaktik Deutsch : Halbjahresschrift für die Didaktik der deutschen Sprache und Literatur* 23 (2018) 45, S. 74-93 - URN: urn:nbn:de:01111-pedocs-203981 - DOI: 10.25656/01:20398

<https://nbn-resolving.org/urn:nbn:de:01111-pedocs-203981>

<https://doi.org/10.25656/01:20398>

in Kooperation mit / in cooperation with:



<http://www.paedagogik.de>

#### Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.  
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

#### Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.  
This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

#### Kontakt / Contact:

**peDOCS**  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Mitglied der

  
Leibniz-Gemeinschaft



Schneider Verlag  
Hohengehren

# Didaktik Deutsch

2018

45

Mit Forschungsbeiträgen von

---

Melanie Bangel, Astrid Müller · Petra Hüttis-Graff,  
Daniel Wirszing · Hans-Georg Müller, Matthias Kepser,  
Stefan Schallenger · Hansjakob Schneider, Miriam Dittmar,  
Eliane Gild, Claudia Schmmellentin

Debatte · Forschungsbeiträge  
Rezensionen · Neuerscheinungen

## **Didaktik Deutsch**

Halbjahresschrift für die Didaktik der  
deutschen Sprache und Literatur

Mitteilungsorgan des Symposions  
Deutschdidaktik e. V.

**<http://www.didaktik-deutsch.de>**

Herausgeber:

Irene Pieper

Thorsten Pohl

Dorothee Wieser

Friederike Kern

23. Jahrgang 2018 – ISSN 1431-4355

Schneider Verlag

Hohengehren GmbH

Wilhelmstr. 13

73666 Baltmannsweiler

Volksbank Plochingen 664309003

(BLZ 611 913 10)

Verantwortlich für den Anzeigenteil:

Ulrich Schneider

Wilhelmstr. 13

73666 Baltmannsweiler

Formatierung: Verlag

Druck: WolfMediePress, D-71404 Korb

Manuskripte bitte einsenden an die  
Herausgeber zu Händen von:

Irene Pieper

**[Irene.Pieper@uni-hildesheim.de](mailto:Irene.Pieper@uni-hildesheim.de)**

Alle Heftbeiträge sind auf Datenträger in einem gängigen Textverarbeitungsformat unter MS-Word oder als rtf-Datei per mail einzureichen. Forschungsbeiträge, Berichte und Debattenbeiträge sind jederzeit erwünscht. Sie sollen an die Herausgeber (s. o.) gesandt werden. Hinweise zur Konzeption der einzelnen Rubriken und zur Textmenge finden sich auf der Homepage der Zeitschrift. Die Forschungsbeiträge werden anonym einer externen Review durch Fachgutachter/innen unterzogen.

Für unverlangt eingesandte Manuskripte wird keine Gewähr geleistet. Die Manuskripte dürfen nicht gleichzeitig anderen Zeitschriften zum Abdruck angeboten werden. – Mit der Annahme des Manuskripts erwirbt der Verlag die ausschließliche Nutzung an allen Verwertungsrechten im Sinne der §§ 15 ff. der Urheberrechtsgesetze, insbesondere das Recht der Vervielfältigung und Verbreitung sowie der Übersetzung. Mit Namen oder Signum gezeichnete Beiträge geben die Meinung des Verfassers, nicht unbedingt auch der Herausgeber wieder.

Die in der Zeitschrift veröffentlichten Beiträge sind urheberrechtlich geschützt. Alle Rechte, insbesondere das der Übersetzung in fremde Sprachen, vorbehalten. Kein Teil dieser Zeitschrift darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form – durch Fotokopie, Mikrofilm oder andere Verfahren – reproduziert oder in eine von Maschinen, insbesondere von Datenverarbeitungsanlagen, verwendbare Sprache übertragen werden. Auch die Rechte der Wiedergabe durch Vortrag, Funk und Fernsehungen, im Magnettonverfahren oder auf ähnlichem Wege bleiben vorbehalten. Fotokopien für den persönlichen und sonstigen eigenen Gebrauch dürfen nur von einzelnen Beiträgen oder Teilen daraus als Einzelkopien hergestellt werden.

Bezug durch Buchhandel und Verlag. – Erscheinungsweise: zweimal jährlich (März, September). Für Mitglieder im Symposion Deutschdidaktik e. V. kostenlos. Bezugspreis jährlich € 24,00 (Vorzugspreis für Studierende € 18,00 bei Vorliegen einer Immatrikulationsbescheinigung) zuzüglich Postgebühren. Preis des Einzelheftes € 14,00. Die Bezugsdauer verlängert sich jeweils um ein Jahr, wenn bis zum 1. Dezember keine Abbestellung erfolgt.

Inhalt

EDITORIAL . . . . .	3
DEBATTE	
<i>Bernard Schneuwly</i>	
Der Praxis ist nie genug: Von der Entstehung der Fachdidaktik aus der Praxis. Gedanken zum Thema „Wie viel Praxis verträgt die Deutschdidaktik?“ . . .	4
<i>Frederike Schmidt</i>	
Die Bezüge in beide Richtungen denken: Plädoyer für produktive Brückenschläge zwischen Fachdidaktik und Praxis . . . . .	11
<i>Björn Stövesand</i>	
Wer schaut wie auf die Praxis? Potenziale einer vermittelnden Perspektive für die Deutschdidaktik . . . . .	18
<i>Thorsten Pohl</i>	
„Die Praxis“ einer erwerbs- und vermittlungsbezogenen Wissenschaft – Eine nachdenkliche Replik . . . . .	22
FORSCHUNGSBEITRÄGE	
<i>Melanie Bangel &amp; Astrid Müller</i>	
Strukturorientiertes Rechtschreiblernen. Ergebnisse einer Interventionsstudie zur Wortschreibung in Klasse 5 mit Blick auf schwache Lerner/ -innen . . . . .	29
<i>Petra Hüttis-Graff &amp; Daniel Wirsching</i>	
Anfänge des Rechtschreiblernens 1994 und 2014 – Ergebnisse einer diachronen Longitudinalstudie in Hamburg . . . . .	50
<i>Hans-Georg Müller, Matthis Kepser &amp; Stefan Schallenberger</i>	
Getrennt- und Zusammenschreibung – ein konsistentes Konstrukt? Erste Ergebnisse einer empirischen Studie . . . . .	74
<i>Hansjakob Schneider, Miriam Dittmar, Eliane Gilg &amp; Claudia Schmellentin</i>	
Textseitige Maßnahmen zur Unterstützung des Leseverstehens im Biologieunterricht . . . . .	94

## REZENSION

*Julia Ricart Brede*

„Sprache und Fach“ oder „Fach und Sprache“? Zwei Bände mit unterschiedlichem Fokus in einer interdisziplinären Debatte . . . . . 117

Neuerscheinungen . . . . . 123

Hans-Georg Müller, Matthis Kepser & Stefan Schallenberger

## GETRENNT- UND ZUSAMMENSCHREIBUNG – EIN KONSISTENTES KONSTRUKT? ERSTE ERGEBNISSE EINER EMPIRISCHEN STUDIE

### *Zusammenfassung*

Das Forschungsprojekt Orthografieerwerb (FORTE) untersucht den Rechtschreiberwerb durch Analyse von Nutzer/-innendaten eines Online-Rechtschreibtrainers. Ein Forschungsschwerpunkt liegt bei der Getrennt- und Zusammenschreibung, die bis dato kaum gesondert didaktisch modelliert und wenig empirisch erforscht worden ist. Erste Ergebnisse einer differenzierten Datenerhebung anhand von ca. 9.000 Versuchspersonen und 180 Testsätzen zeigen, dass die Annahme eines umfassenden Kompetenzkonstrukts „Getrennt- und Zusammenschreibung“ nicht in Frage gestellt werden muss. Das Datenmaterial wird bezüglich der Aufgabenschwierigkeit eingeschätzt. Dies geschieht unter dem Aspekt der zutage tretenden grammatischen, prosodischen, semantischen und syntaktischen Merkmale der getrennt- bzw. zusammenzuschreibenden Einheiten. Untersucht wird es zudem hinsichtlich des Zusammenschreibverhaltens der Probanden sowie verschiedener Persönlichkeitsvariablen wie Klassenstufe, Geschlecht und Schulart. Die Ergebnisse liefern erste Ansätze für ein umfassenderes didaktisches Erwerbsmodell.

### *Abstract*

The research project FORTE explores the acquisition of orthography competences by analysis of user data of an online orthography trainer. One focus of research is on the separate or compound spelling, which has hardly been didactically modeled until now and has not been explored empirically in depth. First results of a differentiated data collection with about 9,000 test subjects and 180 test sentences show that the assumption of a comprehensive competence construct „separate or compound spelling“ has not to be rejected. The data material is estimated in terms of task difficulty. This will be done under the focus of grammatical, prosodic, semantic and syntactic features of the units to be written separated or compound. It is also being explored with regard to the writing behavior of the subjects as well as various personality variables such as grade level, gender and school type. The results provide first approaches for a more comprehensive didactic acquisition model.

## Einleitung

Das Forschungsprojekt Orthografieerwerb untersucht den Rechtschreiberwerb empirisch und nutzt als Basis die Nutzerdaten des Online-Rechtschreibtrainers Orthografietrainer.net, auf dem Schüler und Schülerinnen sowie angemeldete Gäste ihre Rechtschreibung trainieren. Das Übungsmaterial besteht in der Regel aus kontextunabhängigen Sätzen, zu denen die Nutzer/-innen orthografisch relevante Entscheidungen treffen müssen. Zusätzlich besteht die Möglichkeit, sogenannte Kompetenztests zu absolvieren, bei denen unterschiedliche orthografische Probleme getestet werden. Eine Rückmeldung über die Normrichtigkeit der jeweiligen Entscheidung der Nutzer erfolgt stets am Ende des Kompetenztests. Bisher haben sich dort über 400.000 Nutzer/-innen aus allen Schularten, Altersstufen und deutschsprachigen Ländern registrieren lassen.

Orthografiedidaktisch interessant sind Fehlerrückmeldungen und Beschreibungen von Lernfortschritten in Kombination mit Informationen zu Alter, Geschlecht, Schulart und Region. Weitere personenbezogene Daten wie Bildungshintergrund, Milieu u. v. m. werden seit Herbst 2016 von den Nutzer/-innen erbeten; sie liegen bisher von über 9.000 Anwender/-innen vor. Innerhalb des FORTE-Projekts sind für die kommenden Jahre verschiedene Untersuchungen geplant, so die Testung von Hypothesen zu theoretischen Annahmen des Rechtschreiberwerbs, die Untersuchung der Wirksamkeit von Hilfestellungen sowie der Effizienz unterschiedlicher Übungsmethoden in Bezug auf den Kompetenzerwerb (Forschung). Nicht zuletzt dienen die erhobenen Daten dazu, die Trainingsplattform selbst weiter zu verbessern (Entwicklung).

Was FORTE im Forschungszusammenhang leisten kann, soll im Folgenden am Problembereich „Getrennt- und Zusammenschreibung“ in ersten Ansätzen dargestellt werden.

## Norm, System und Didaktik der Getrennt- und Zusammenschreibung (GZS)

Wegweisend für die GZS ist in systematischer Hinsicht die Unterscheidung zweier Prinzipien, nämlich eines morphologischen oder Wortbildungsprinzips und eines syntaktischen oder „Relationsprinzips“, wie Fuhrhop (2015b: 107) herausstellt.<sup>1</sup> Viele Komposita oder kompositaähnliche Verbindungen können in Abhängigkeit vom Kontext zusammen- oder getrennt geschrieben werden (vgl. ebd.). Solche Schreibungen seien, so Fuhrhop, den meisten Schreibern jedoch intuitiv völlig klar (vgl. ebd.). Insgesamt erweist sich die GZS als relativ einfach, wenn sich die Schreibung entweder am Wortbildungsprinzip oder dem Relationsprinzip orientiert. Erst

<sup>1</sup> Auch Eisenberg (2017: 31) verweist mit dem „Wortbildungsprinzip, vorgeschlagen in Jacobs 2005“ (ebd.) und dem Relationsprinzip nach Fuhrhop auf diese beiden „einfachen Prinzipien“ (ebd.), mithilfe derer man „nicht alles, aber sehr vieles“ (ebd.) erreichen könne.

wenn beide Prinzipien relevant werden und über die Dominanz eines der beiden Prinzipien hinsichtlich der GZS entschieden werden muss, wird ein sogenannter Peripheriebereich diskutiert. Daher lassen sich nach Fuhrhop in diesem Spektrum neben einem unproblematischen Kernbereich ein schwierigerer Teil des Kernbereichs und schließlich der Peripheriebereich unterscheiden (vgl. ebd.: 115), was anhand konkreter Beispiele ausführlich in Fuhrhop 2015a und 2015b verdeutlicht wird. Im Rahmen des von ihr gegebenen Forschungsüberblicks (Fuhrhop 2015b) wird auch die etwas abweichende Konzeption von Jacobs (2005) dargestellt und diskutiert, dessen noch vor der zweiten Reform erschienene Monographie „von vornherein das Gesamtsystem der Getrennt- und Zusammenschreibung darzustellen [versucht]“ (Fuhrhop (2015b: 110).

Hinsichtlich einer Didaktik der GZS bemerkt Birgit Mesch: „Bislang existiert weder eine systematisch konzipierte noch eine speziell auf die GZS zugeschnittene Didaktik, welche die [...] Anstrengungen und Leistungen von Lerner/innen [...] adäquat berücksichtigt“ (Mesch 2015: 278). Sie fährt fort: „Allgemeine Beiträge zur Rechtschreibdidaktik behandeln die GZS gar nicht [...], nur Einzelfälle bzw. den Peripheriebereich [...], erst in der Sekundarstufe [...] oder unter falscher Prämisse, nämlich auf Wortebene“ (ebd.: 278). Grundlegend wird der Hinweis gegeben, dass die amtliche Regelung von 2006 nicht ausreichend systematisch fundiert (ebd.: 268), sondern stellenweise auch „kasuistisch“ geregelt sei (ebd.: 269). Was bei Augst/Dehn (2007: 143) mit „Intuition“ erfasst werde,<sup>2</sup> gelte in der Erwerbstheorie als Indiz dafür, „dass sich Lerner/innen die Schrift eigenaktiv, hypothesenbildend und (re-)konstruktiv aneignen“ (ebd.: 278). Blatt (2015) unternimmt ausgehend von den systematischen Überlegungen Fuhrhops schließlich den Versuch einer modellhaften Unterrichtskonzeption zur GZS für die Sekundarstufe I. Diese verbindet im Kernbereich eine systematische Arbeit zu grammatischen Strukturen mit sprachlichem und literarischem Lernen in Form von entdeckendem und selbstkontrolliertem Vorgehen und gibt praktische Hilfestellungen für den Peripheriebereich (vgl. Blatt 2015: 572).

Die bisher vorliegenden empirischen Studien zur orthografischen Kompetenz bezüglich GZS finden sich bei Nottbusch/Jonischkait (2015: 173–175) zusammengefasst. Sie unterscheiden dabei „Korpusstudien“ (ebd.: 173), die u. a. auf Grundlage von „Diktat, Aufsatz, ggf. Facharbeit“, „Aufsätzen von Abiturient/innen“ (ebd.) oder „freier Textproduktion“ (ebd.: 173 f.) entstanden sind, und „Experimente“ im Hinblick auf die „Rezeption“ hinsichtlich der „Leseleistung“ (ebd.) oder die „Produktion“ (ebd.: 175) mit Augenmerk auf „verschriftete Bilder-

<sup>2</sup> Sie formulieren unter der Überschrift „Unterrichtliche Hinweise zur Getrennt- und Zusammenschreibung und zum Bindestrich“ eingangs: „Die GuZ ist als unterrichtlicher Gegenstand nur etwas für die höheren Klassen, da neben der Prüfung von Intonation und Wortbedeutung auch ein gerüttelt Maß an grammatischen Kenntnissen notwendig ist, um die Zusammenhänge bewusst zu durchschauen. Vieles lernen Kinder intuitiv, so z. B. die Zusammenschreibung von untrennbaren Verben, von Partikeln + Verb, von Substantiven, und es ist daher vielleicht besser, gar nicht daran zu rühren“ (Augst/Dehn (2007: 143).



geschichten“ (ebd.). Die angeführten Studien belegen, dass der „GZS [...] aus empirischer Sicht bisher wenig Aufmerksamkeit geschenkt worden [ist]“ (ebd.). Viele Fehlerstatistiken, die im Vorfeld der Rechtschreibreform entstanden sind (etwa Riehme/Heidrich 1970, Menzel 1985, Zimmermann/Heckel 1986 u. a.), nennen darüber hinaus prozentuale Fehleranteile, die aber auf keine Grundgesamtheit bezogen werden. In zahlreichen Studien (etwa Valtin et al. 2003) wird die GZS ausgeklammert, was bereits Mesch (2015: 274) moniert. Die wenigen existierenden Daten deuten darauf hin, dass Zusammensetzungen im verbalen Bereich wesentlich fehleranfälliger sind als Zusammensetzungen im nominalen Bereich (vgl. nochmals Nottbusch/Jonischkait 2015: 175). Mesch konstatiert ferner mit Bezug auf frühere Forschungsergebnisse, dass falsche Zusammenschreibung signifikant häufiger bei türkischen Schülern als bei deutschen Schülern auftritt (vgl. Mesch 2015: 275). Vor diesem Hintergrund will die hier vorgestellte FORTE-Untersuchung einen weiteren empirischen Beitrag zum System und Erwerb der GZS leisten. Eine wichtige Forschungsfrage wird dabei sein, ob sich die Annahme einer eindimensionalen Kompetenz zur Getrennt- und Zusammenschreibung überhaupt aufrechterhalten lässt.

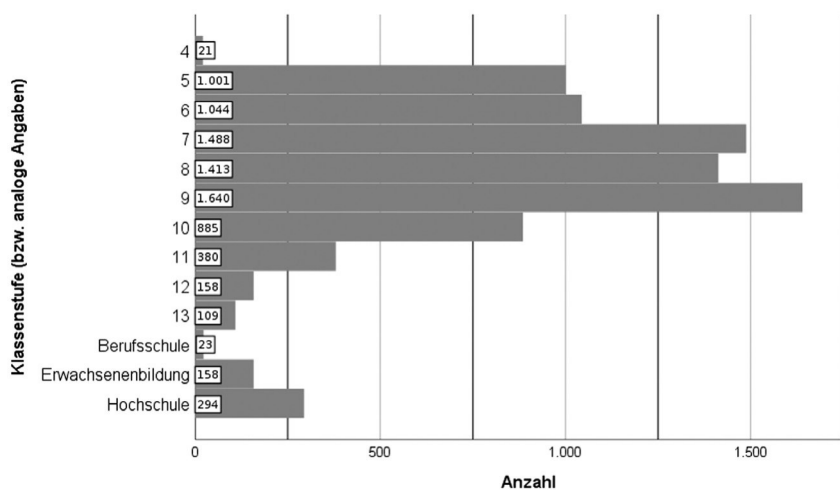


Abb. 1 Verteilung der Untersuchungskohorte nach Klassenstufen der Grund- und Sekundarstufe (bzw. analogen Angaben für andere Bildungsgänge)

## Eine empirische Studie zur GZS

Die den folgenden Ausführungen zugrunde liegenden Daten basieren auf den Antworten von 8.985 Probanden auf insgesamt 180 Aufgaben zur Getrennt- und Zusammenschreibung. Angemeldet haben sich die Probanden überwiegend auf Initiative ihrer Lehrkräfte, um die Rechtschreibleistung zu trainieren. 6.350 der Versuchspersonen stammen aus Deutschland, 1.732 aus Österreich und 449 aus der Schweiz. Beim Großteil der Versuchspopulation handelt es sich um Schüler/-innen

der Sekundarstufe unterschiedlicher Schulformen; über 80 % der Probanden haben sich den Klassenstufen 5–10 zugeordnet (vgl. Abb. 1). Für sie spielt also der fortgeschrittene Rechtschreiberwerb die entscheidende Rolle.

Die Aufgaben der Kompetenztests zur GZS decken 30 zentrale Anwendungsfälle der amtlichen Regelungen mit jeweils sechs Einzelaufgaben ab. Aus methodischen Gründen wurde dabei der Fokus auf komplexere orthografische Probleme gelegt: So verursachte etwa die Schreibung einfacher Nominalkomposita bereits bei jüngeren Probandengruppen kaum mehr Fehler.<sup>3</sup> Diese Aufgaben wurden daher aus den Tests entfernt.

Soweit möglich, wurden die Anwendungsfälle bewusst parallel konstruiert, also bspw. Nomen-Verb-Verbindungen sowohl in Getrennt- als auch in Zusammenschreibung. Dabei wurde das grundlegende Strukturelement der amtlichen Regelungen, die Wortartenkategorie, als Basis der Aufgabenbildung übernommen. Gut drei Viertel der Anwendungsfälle prüft die Getrennt- und Zusammenschreibung verschiedener Kombinationen von Inhaltswörtern (Nomina, Adjektive, Verben, Adverbien, Partizipien), etwa ein Viertel deckt Aufgaben zur Schreibung komplexer Funktionswörter (Konjunktionen, Präpositionen, Pronomina, Partikel) ab.

Um die Modellqualität aufrechtzuerhalten, wurden in die Untersuchung nur die Daten von Probanden aufgenommen, die mindestens ein Drittel der Testaufgaben bearbeitet hatten. Aus den Antwortpatterns wurde mithilfe einer dichotomen Rasch-Skalierung (vgl. Rost 2004: 115–132, Thomé/Gomolka 2007: 143) für jede Versuchsperson ein Kompetenzparameter und für jede Aufgabe ein Schwierigkeitsparameter geschätzt. Die Skalierung mithilfe des Rasch-Modells gewährleistet eine grundlegende Prüfung der fachdidaktisch motivierten Annahme, dass es sich bei der Anwendung der Getrennt- und Zusammenschreibung um ein konsistentes, empirisch fundiertes Kompetenzkonstrukt handelt, in dem sich aus der Anzahl gelöster Aufgaben sinnvoll auf die Kompetenz zur normgerechten GZS der Person und aus der Häufigkeit der normgerechten Bearbeitung auf die Schwierigkeit der jeweiligen Aufgabe schließen lässt (für Einzelheiten vgl. Müller 2016: 249–257). Die Begutachtung der Fit-Statistiken zeigte, dass die Daten den Annahmen des Rasch-Modells sehr gut entsprachen. Von einem Ausreißer abgesehen<sup>4</sup> hatten alle Items eine mindestens akzeptable und meist gute Trennschärfe (unweighted fit:  $0,89 \leq CI \leq 1,11$ ) und indizierten damit, dass die Annahme eines globalen Kompetenzkonstrukts Getrennt- und Zusammenschreibung zunächst nicht infrage gestellt werden muss und weiter aufrechterhalten werden kann. Vorläufig nicht untersucht wurde ein Vergleich des einfachen Rasch-Modells mit mehrdimensionalen Skalierungsverfahren, deren Diskussion den Umfang dieses Beitrages sprengen würde.

<sup>3</sup> Die sechs zunächst in den Pilottest aufgenommenen nominalen Komposita zeigten durchgängig eine Fehlerrate von weniger als 0,1 %. Das leichteste Item („Hundehütte“) wurde zu 100 % richtig gelöst.

<sup>4</sup> Es handelte sich um die Nominalisierung „das Dabeisein“, die nur einmal falsch geschrieben wurde.

## Aufgabeneffekte

Die Aufgaben des Testmaterials wurden anhand verschiedener grammatischer Merkmale charakterisiert, die sich neben den Strukturmerkmalen der amtlichen Regelungen insbesondere aus den hypothetischen Annahmen von Jakobs (2005: 128–141) ergeben. Spezifiziert wurde dabei die Wortartenkategorie von Erst- und Zweitglied, die Betonung sowie verschiedene grammatische Eigenschaften, aus denen sich auf die semantisch-konzeptuelle (Un-)Abhängigkeit der beiden Glieder schließen lässt, so etwa die topologische Trennbarkeit beider Glieder, die freie Austauschbarkeit des Erstgliedes, die separate Vorfeldfähigkeit sowie verschiedene Operationalisierungsmittel, mit denen auf die Anwendbarkeit des Kompositionalitätsprinzips geschlossen werden kann. Um ferner Effekte der Vertrautheit mit dem Schriftbild zu erheben, wurde die Häufigkeit der Wortverbindung anhand des Wortschatzkorpus der Universität Leipzig<sup>5</sup> ermittelt. Ziel der grammatischen Charakterisierung der Items war es, das Verhalten der Versuchspopulation in Bezug auf die Aufgaben statistisch zu modellieren, also die Schwierigkeit einer Aufgabe als ein Resultat ihrer grammatischen Eigenschaften zu verstehen.

## Zur Modellierung der Aufgabenschwierigkeit

Das dichotome Rasch-Modell liefert einen Messwert der Aufgabenschwierigkeit, dessen statistische Kennwerte in Tabelle 1 wiedergegeben sind. Ein hoher Messwert indiziert eine hohe Schwierigkeit und damit eine niedrige Lösungshäufigkeit durch die Probanden. Da die mittlere Itemschwierigkeit auf Null normiert wurde (constraints by item), zeigen die Daten an, dass der Schwierigkeitsparameter nach oben stärker streut als nach unten.

Tab. 1: Statistische Kennwerte der Aufgabenschwierigkeit

	N	Minimum	Maximum	Mittelwert	Std.abweichung
IP	179	- 2,778	3,534	,0945	1,2655

Eine erste Begutachtung zeigt, dass sich unter den besonders leichten Items viele Verbindungen mit Adjektiven und Verben als Zweitglied finden (*eiskalt, wortreich, sinnvoll, lecker riechen, reden hören, warm werden* etc.). Unter den besonders schwierigen Items fallen insbesondere mehrteilige Funktionswörter wie Pronomina und Pronominaladverbien ins Auge (*sonst wer, infolgedessen*), ferner Konjunktionen (*sooft*), Interjektionen (*na ja, ach so*) sowie Verbverbindungen, die nicht dem Kompositionalitätsprinzip entsprechen (*vorliebnehmen, dazwischengehen, heiligsprechen* etc.).

<sup>5</sup> <http://wortschatz.uni-leipzig.de/>, Recherchedatum 12./13. 10. 2016.

Um zu prüfen, inwieweit sich die Schwierigkeiten der normrichtigen Getrennt- und Zusammenschreibung aus den grammatischen Merkmalen der Items erklären lassen, wurde ein lineares Regressionsmodell<sup>6</sup> geschätzt. Es erwies sich allerdings als enttäuschend unspezifisch: Die Schätzung kann lediglich 23,9% der Varianz des Schwierigkeitsparameters aufklären, was in Anbetracht paralleler Datenmodelle<sup>7</sup> bemerkenswert wenig ist.

Zur Behebung dieses Mankos wurden verschiedene Parallelmodelle geschätzt, bei denen insbesondere die separate Betrachtung von getrennt und zusammenzuschreibenden Items einige bemerkenswerte Effekte hervorrief. So stieg der Prozentsatz aufgeklärter Varianz bei singulärer Betrachtung der Items mit Getrenntschreibung auf 67,0% (angepasstes  $R^2 = .670$ ,  $p < .001$ ), bei der Zusammenschreibung immerhin noch auf 39,2% (angepasstes  $R^2 = .392$ ,  $p < .001$ ). Auf eine gesonderte Darstellung soll an dieser Stelle dennoch verzichtet werden, da die Aufteilung des Korpus nach Getrennt- und Zusammenschreibung einige unerwünschte orthografiedidaktische Folgen mit sich bringt.

Gleichwohl zeigt die unterschiedliche Aufklärungsrate für getrennt bzw. zusammengeschriebene Items, dass die gewählten grammatischen Einflussvariablen besser geeignet sind, die Getrennt- als die Zusammenschreibung zu modellieren. Dieser Befund deckt sich mit dem graphematischen Prinzip, dass Getrenntschreibung den Normalfall darstellt, während Zusammenschreibung durch ganz unterschiedliche grammatische Prozesse verursacht werden kann, was den höheren Anteil unaufgeklärter Varianz in diesem Bereich erklärt.

Tabelle 2 gibt an, welche der grammatischen Merkmale einen signifikanten Einfluss<sup>8</sup> auf die Schwierigkeit zur Getrennt- und Zusammenschreibung haben, wobei ein positiver Wert die Steigerung der Schwierigkeit anzeigt, ein negativer die Senkung. Einige, insbesondere funktionale Wortartenkategorien zeigen eine signifi-

<sup>6</sup> Die Auswahl der unabhängigen Variablen erfolgte anhand der orthografiethoretischen Diskussion, die weiter oben erwähnt wurde. Die simultane Betrachtung aller Einflussfaktoren entspricht der vorläufigen Annahme, dass alle grammatischen Eigenschaften gleichzeitig, wenn auch nicht mit gleichem Gewicht zur Einschätzung der GZS herangezogen werden. Ob diese Annahme gerechtfertigt ist, kann nur durch Schätzung alternativer Modelle getestet werden, die späteren Untersuchungen vorbehalten bleibt.

<sup>7</sup> In Müller (2006: 128 f.) wurden mit analoger Methodik für die Kommasetzung Aufklärungsquoten von 78,7% erreicht, in Müller (2016: 286) für die Groß- und Kleinschreibung immerhin noch 51,6%.

<sup>8</sup> Der Konvention gemäß wurden auch Variablen im Modell nicht ausgeschlossen, deren Signifikanzniveau  $.05 < p < .10$  beträgt, also schwach oberhalb der Signifikanzschwelle liegt. Die entsprechenden Variablen sollten allerdings nicht oder nur sehr vorsichtig interpretiert werden. Auch das Problem der Multikollinearität muss für die Interpretation der einzelnen Modellparameter berücksichtigt werden, da die analysierten grammatischen Eigenschaften strukturbedingt teils hoch korreliert sind. Die Folge davon ist, dass es schwer ist, die Höhe des tatsächlichen Einflusses jeder einzelnen Variable präzise einzuschätzen, was sich nicht zuletzt in den hohen Standardfehlern niederschlägt. Für die Interpretation bedeutet das insbesondere, dass die tatsächliche Höhe der Koeffizienten nur annäherungsweise eingeschätzt werden kann und geringe Unterschiede nicht zu stark interpretiert werden dürfen.

kante Steigerung der Aufgabenschwierigkeit an, die teils knapp unter einer Standardabweichung, teils deutlich darüber liegt, was den Augenschein der ersten Begutachtung bestätigt (s. o.).

Tab. 2: Koeffizienten Regressionsanalyse, abhängige Variable: Itemparameter

Koeffizienten (nicht standardisiert)				
Einflussvariable	B	Std.fehler	F	Sig.
(Konstante)	-,523	,461	-1,136	,258
Betonung auf dem Erstglied	-,693	,387	-1,792	,075
Betonung auf dem Zweitglied	-,705	,308	-2,292	,023
Häufigkeitsklasse	,076	,020	3,746	,000
Erstglied allein vorfeldfähig	,780	,295	2,639	,009
Einfügeprobe gelingt	-,700	,309	-2,265	,025
Transp_1_Konst	-,352	,208	-1,695	,092
2_ist_Adverb	2,197	,453	4,847	,000
1_ist_Partizip1	1,787	,508	3,521	,001
2_ist_Substantiv	,747	,301	2,481	,014
1_ist_Verb	-1,027	,445	-2,307	,022
2_ist_Pronomen	1,067	,304	3,510	,001
2_ist_Partikel	1,125	,379	2,968	,003

Der einzig negative Koeffizient in den Wortartenkategorien tritt bei Verbindungen mit Verben als Erstglied auf, vor allem bei Verb-Verb-Kombinationen (*bleiben lassen, tanzen gehen, sprechen üben* etc.). Dagegen zeigten Verbindungen mit Zweitglied Substantiv eine erschwerende Tendenz, die insbesondere von verschiedenen Denominalisierungen (*zuhauf, jederzeit, zurzeit* etc.) verursacht wurde. Die Befunde scheinen den Aussagen von Nottbusch/Jonischkait (2015: 175) zu widersprechen, dass Fehler im verbalen Bereich häufiger auftreten als im nominalen. Tatsächlich dürfte es sich um eine Frage der Itemauswahl handeln, da einfache nominale Komposita aus statistischen Gründen aus der Untersuchung ausgeschlossen worden waren, während Denominalisierungen bei Nottbusch/Jonischkait keine Rolle spielten (vgl. ebd.).

Die Betonungsmuster zeigen beide einen erleichternden Einfluss, der bei der „Betonung auf dem Erstglied“ knapp das Signifikanzniveau verfehlt. Dass beide Betonungsmuster einen erleichternden, aber statistisch nur bedingt gesicherten Effekt haben, ist auf eine Reihe von Items zurückzuführen, bei denen beide Glieder betonbar sind (*sicher vermeiden, gut sprechen, lecker riechen* etc.). Schwerere

Items weisen häufiger nur ein mögliches Betonungsmuster auf, das aber sehr ähnlich auf Erst- und Zweitglied verteilt ist und damit zur statistischen Unsicherheit der Variable beiträgt.

Ein eindeutiger Effekt ist für die Häufigkeitsklasse der Wortverbindung nachweisbar, dessen positiver Koeffizient zeigt, dass seltenere Kombinationen signifikant schwieriger sind. Der Effekt entspricht damit den Erwartungen.

Unter den semantischen Kriterien zeigt das Gelingen der Einfügeprobe, also der Möglichkeit, zwischen die beiden Konstituenten ein weiteres Syntagma zu platzieren, erwartungsgemäß eine Erleichterung an. Auch die semantische Transparenz des Erstgliedes scheint die normkonforme Schreibung zu erleichtern, was allerdings mit  $p = ,092$  nicht hinreichend gegen den Zufall abgesichert ist. Der Grund hierfür dürfte darin liegen, dass semantische Transparenz tendenziell Getrennschreibung nahelegt, ihr Fehlen aber ebenfalls als erleichterndes Signal – nämlich für Zusammenschreibung – wirken kann.<sup>9</sup>

Bemerkenswert ist der erschwerend wirkende Einfluss der Vorfeldfähigkeit, der den Erwartungen widerspricht, da die Vorfeldprobe ein vergleichsweise sicherer Indikator für Getrennschreibung ist. Tatsächlich weisen fast alle Items mit allein vorfeldfähigem Erstglied ein verbales Zweitglied auf (Auto fahren, schwerfallen, voranbringen etc.). Ihre normrichtige Getrennt- und Zusammenschreibung fiel der Versuchspopulation tendenziell schwer, nicht zuletzt freilich deshalb, weil die Anwendungsbeispiele bewusst so gewählt worden waren, dass die schwierige Grenze zwischen komplexem Wort und Syntagma im Modell möglichst scharf abgebildet werden konnte. Besser als die oben beschriebenen Daten zeigen diese Fälle, dass der verbale Bereich analog zu Nottbusch/Jonischkait (2015: 175) tendenziell größere Schwierigkeiten bereitet als der nominale.

Insgesamt erweist sich die Modellierung der Aufgabenschwierigkeit als durchgängig plausibel und weitgehend erwartungskonform. Sie scheint die Annahme Bredels (2006: 147) zu erhärten, dass Rechtschreibernende während des Normerwerbs mit unterschiedlichen phonologischen, morphologischen, grammatischen und semantischen Eigenschaften experimentieren, um auf die normrichtige Schreibung zu schließen. Fraglich bleibt allerdings, welche Eigenschaft im Laufe des Erwerbsprozesses welche Rolle spielt. Die Beantwortung dieser Frage setzt ein Kompetenzstufenmodell der GZS voraus, das den Rahmen dieses Artikels sprengen würde.

### Zur Modellierung des Zusammenschreibverhaltens

Unabhängig von der Frage der Normrichtigkeit kann anhand des Datenmaterials untersucht werden, inwiefern die erhobenen grammatischen Merkmale einen Einfluss auf die Tendenz der Versuchspopulation zeigen, die entsprechende Textstelle als komplexes Wort zu interpretieren und folglich zusammenzuschreiben. Dazu

<sup>9</sup> Bei der Einfügeprobe ist dieses Problem weit weniger gegeben, da diese viel stärker mit Getrennschreibung korreliert ( $r = ,648$ ,  $p < ,001$ ).

wurde mittels einer weiteren Regressionsanalyse untersucht, inwiefern sich die prozentuale Häufigkeit der Zusammenschreibung als Funktion seiner grammatischen Merkmale modellieren lässt. Unberücksichtigt blieb dabei, ob die Entscheidung über die Getrennt- bzw. Zusammenschreibung normkonform war oder nicht (zur methodischen Diskussion vgl. letzten Abschnitt).

Tab. 3: Modellzusammenfassung Regressionsanalyse, abhängige Variable: Itemparameter

Modell	R	R <sup>2</sup>	Angepasstes R <sup>2</sup>	Std.fehler der Schätzung
2	,860	,739	,715	,173620

Bereits die Betrachtung der Modellübersicht (Tab. 3) zeigt eine massiv gestiegene Aufklärungsrate, die mit 71,5 % (angepasstes  $R^2 = .715$ ,  $p < .001$ ) beinahe dreimal so hoch ist wie im letzten Modell. Die gewählten grammatischen Einflussfaktoren sind also durchaus in der Lage, das Verhalten der Versuchspopulation maßgeblich aufzuklären, sofern dafür von der Normrichtigkeit der Entscheidung abgesehen wird. Mit anderen Worten: Unregelmäßig ist nicht das graphematische Verhalten der Probanden, sondern dessen Verhältnis zur gültigen Norm. Dieser Umstand ist auf den ersten Blick überraschend, deckt sich allerdings mit den Befunden paralleler statistischer Erhebungen und lässt sich durch weitere Überlegungen plausibilisieren. So sind bspw. unterschiedliche orthografische Prüfstrategien geläufig, etwa die Betonungsprobe, die bereits bei Jakobs (2005: 135) diskutiert wird und die bei den Items des Korpus sowohl zu normkonformer wie auch zu normwidriger Schreibung führen kann. Die Betonungsvariablen konnten daher die Aufgabenschwierigkeit kaum erklären. Das ist bei der Modellierung der Zusammenschreibungshäufigkeit durchaus anders, wie gleich näher erläutert wird.

Tabelle 4 (Koeffizienten) zeigt die signifikanten Einflussvariablen des Regressionsmodells und bestätigt den Befund der weitaus besseren Passung. Die Variablen liegen fast durchgängig auf hohem, oft höchstem Signifikanzniveau.

Tab. 4: Koeffizienten Regressionsanalyse, abhängige Variable: Itemparameter

Koeffizienten (nicht standardisiert)				
Einflussvariable	B	Std.fehler	T	Sig.
(Konstante)	,426	,089	4,809	,000
Betonung auf dem Erstglied	,138	,070	1,975	,050
Betonung auf dem Zweitglied	-,131	,048	-2,713	,007
Erstglied allein vorfeldfähig	-,205	,050	-4,115	,000
Einfügeprobe gelingt	-,321	,049	-6,515	,000
Polysemie Homonymie	,147	,045	3,254	,001
Transp_1_Konst	,088	,035	2,513	,013
1_ist_Adverb	,126	,046	2,741	,007
1_ist_Substantiv	,100	,044	2,263	,025
2_ist_Substantiv	,153	,055	2,769	,006
2_ist_Adjektiv	,166	,041	4,007	,000
1_ist_Verb	-,245	,076	-3,251	,001
1_ist_Partikel	,153	,072	2,115	,036
1_ist_Präposition	,322	,068	4,719	,000
1_ist_Konjunktion	,477	,089	5,360	,000
2_ist_Konjunktion	,320	,189	1,696	,092

Die Variablen zur Betonung weisen mit beinahe identischem Betrag in unterschiedliche Richtungen: Betonung des Erstgliedes steigert die Tendenz zur Zusammenschreibung leicht, Betonung des Zweitgliedes senkt sie leicht ab. In diesem Befund spiegelt sich die implizite Kenntnis und weithin verbreitete Prüfstrategie wider, dass die Betonung des Erstgliedes Zusammenschreibung signalisiert, was im Korpus für viele Kombinationen von Inhaltswörtern eine adäquate Lösung ermöglichte, bei einigen Funktionswörtern allerdings auch in die Irre führte (*so oft* vs. *sooft* etc.). Die Betonungsprobe könnte damit die hohen Schwierigkeiten bei der Schreibung komplexer Funktionswörter mit verursacht haben.

Auch die Vorfeldfähigkeit des Erstgliedes und die Einfügeprobe, die beide die Tendenz zur Getrennschreibung verstärken, können als Niederschlag mehr oder weniger impliziter Prüfstrategien betrachtet werden, wobei die Variable „Einfügeprobe gelingt“ mit  $B = -,321$  einen noch stärkeren Effekt zeigt.

Beide Variablen sind zwar kein Beleg für die Anwendung der entsprechenden Strategie, wohl aber dafür, dass die Versuchspersonen über ein mehr oder weniger tief ausgeprägtes „Gespür“ für die semantische und syntaktische Unabhängigkeit des



entsprechenden Textbestandteils besitzen, welche sich in den Proben niederschlägt und ihrerseits systematisch mit der orthografischen Norm zusammenhängt.

In ähnlicher Weise muss wohl der positive Einfluss der Variable „Polysemie/Homonymie“ verstanden werden. Die Variable indiziert Mehrdeutigkeiten in der Kombination der Konstituenten (*wieder holen* vs. *wiederholen* etc.). Dass die Versuchspopulation bei diesen Items tendenziell mit verstärkter Zusammenschreibung reagiert, ist möglicherweise ein Resultat visueller Habituation i. S. einer Wiedererkennung eines existierenden Wortbildes.

Der geringe, aber doch signifikante Effekt der Variable „Semantische Transparenz des Erstgliedes“ rührt wahrscheinlich von einem Charakteristikum des Korpus her, in dem bei mehreren zusammengeschriebenen Items mit idiosynkratischer Bedeutung die erste Konstituente als semantisch transparentes Resultativ zu verstehen ist (*freisprechen*, *krankschreiben* etc.). Solche Konstruktionen indizieren die Tendenz, dass die übertragene Bedeutung vieler komplexer Wörter erst durch die Kombination mit der zweiten Konstituente entsteht. Sie sind designbedingt im Korpus deutlich häufiger als in nativen Texten.

Unter den Wortartenkategorien des Erstgliedes zeigen sehr viele Variablen einen signifikanten Effekt, sodass vor der Analyse die Charakteristik der abhängigen Variable „Zusammenschreibhäufigkeit“ berücksichtigt werden muss. Diese hat einen Mittelwert von  $m = 0,629$  ( $sd = 0,325$ ), während die Konstante des Regressionsterms mit  $B = 0,426$  deutlich darunter liegt. Die Variablen zur Wortartenkategorie müssen daher, sieht man von der Wirkung aller anderen Variablen ab, einen Wert von mindestens  $B = 0,203$  erreichen, um auch nur in einen mittleren Bereich der Zusammenschreibhäufigkeit zu gelangen. Die große Anzahl positiver Koeffizienten heißt also nicht automatisch, dass sich alle Wortarten positiv auf die Zusammenschreibhäufigkeit auswirken. Vielmehr muss für eine sinnvolle Interpretation das Verhältnis der Variablen zueinander berücksichtigt werden.

Unter diesem Aspekt betrachtet, fällt zunächst der einzig negative Koeffizient bei der Variable „Erstglied Verb“ ( $B = -,245$ ) auf, welcher die starke Tendenz widerspiegelt, Verben von ihren Folgegliedern zu trennen. Da es sich bei den zugehörigen Zweitgliedern fast durchgängig ebenfalls um Verben handelte, deutet die Variable darauf hin, dass die Regel zur Getrennschreibung von Verben in der Versuchspopulation vergleichsweise sicher etabliert ist.

Die übrigen Erstgliedvariablen zeigen bei Inhaltswörtern einen eher kleinen, bei Funktionswörtern einen eher großen Koeffizienten: Den niedrigsten Wert mit  $B = ,100$  haben Substantive, gefolgt von Adverbien ( $B = ,126$ ) und Adjektiven ( $B = ,166$ ). Die Koeffizienten aller drei Kategorien würden für sich genommen noch nicht genügen, den Mittelwert der Zusammenschreibhäufigkeit zu erreichen. Da-

gegen zeigen Präposition<sup>10</sup> ( $B = ,322$ ) und Konjunktion ( $B = ,477$ )<sup>11</sup> deutlich stärkere Effekte und indizieren eine überdurchschnittlich häufige Zusammenschreibung. Damit lässt sich eine Tendenz ausmachen, Inhaltswörter von ihren Folgegliedern getrennt zu schreiben, Funktionswörter hingegen zusammen. Lediglich die Klasse der Partikel fällt mit  $B = ,153$  etwas aus diesem Trend, was sich wohl aus der funktionalen Heterogenität dieser Klasse erklären lässt. Für diese Deutung spricht auch der hohe Standardfehler der Variable.

Unter den Erstgliedvariablen fehlen insbesondere die Inhaltswortkategorien Adjektiv und Partizip I, was statistisch gewissermaßen dem Koeffizienten  $B = 0$  entspricht. Die typischen Wortarten für verbnähe Adverbiale und pränominale Attribute werden folglich eher selten mit dem Folgeglied zusammengeschieden. Wie bei der Getrenntschreibung der Verben drückt sich in diesem indirekten Befund ein relativ solides Grundwissen der Versuchspopulation über typische Satzstrukturen des Deutschen aus.

Unter den Variablen zur Wortart des Zweitgliedes zeigen nur Substantiv und Adjektiv einen statistisch signifikanten und leicht positiven Effekt.<sup>12</sup> Da hier keine weiteren Variablen das Signifikanzniveau erreichen, ist der Effekt trotz seiner nur mittleren Höhe bemerkenswert, zumal die beiden Wortartenkategorien funktional nah verwandt sind, die beinahe denselben Koeffizienten aufweisen ( $B_{\text{Subst.}} = ,153$ ;  $B_{\text{Adj.}} = ,166$ ). Substantive und Adjektive laden also in etwa gleicher Weise dazu ein, ein davorstehendes Lexem als morphologische Erweiterung zu interpretieren und zusammenzuschreiben.

Insgesamt zeigt die Analyse des Zusammenschreibverhaltens eine bemerkenswerte Sensibilität der Versuchspopulation für sprachliche Regularitäten, die im Zusammenhang mit der Getrennt- und Zusammenschreibung relevant sind. Wiederum wird damit die Annahme Bredels (2006: 147) bestärkt, dass der Erwerb der GZS-Norm eine Art grammatisches Bootstrapping voraussetzt: Die Ausprägung der Variablen offenbart eine Systematik im orthografischen Verhalten der Probanden, die von hoher fachdidaktischer Relevanz ist, weil sie Hinweise darauf gibt, nach welchen Kriterien Schreiblernende im Stadium des fortgeschrittenen Orthografieerwerbs auf die richtige Schreibung einer Textstelle zu schließen suchen.

Dass sich die Zusammenschreibhäufigkeit statistisch erheblich besser modellieren lässt als die Rechtschreibschwierigkeit, bedeutet: Orthografische Fehler sind keine zufälligen Abweichungen vom gültigen orthografischen System, sondern Resultat eines hoch systematischen Verhaltens der Versuchspersonen – mit der Einschrän-

<sup>10</sup> Der Begriff ist hier streng genommen nicht angemessen, da es sich in den allermeisten Fällen eigentlich um mit Präpositionen homonyme Verbpartikel handelt (vgl. Eisenberg 2017: 70 f.). Eine Kategorisierung als Präposition war hier gleichwohl sinnvoll, da Präposition und homonyme Partikel in systematischer Korrespondenz zueinander stehen. Analog wurde im Korpus mit anderen Funktionswortkategorien gearbeitet.

<sup>11</sup> Vgl. Anm. 10.

<sup>12</sup> Der Koeffizient des Zweitgliedes Konjunktion ist mit  $B = ,320$  zwar wesentlich höher, weist aber auch einen sehr hohen Standardfehler auf und verfehlt mit  $p = ,092$  das Signifikanzniveau.

kung, dass die darin zum Ausdruck kommende Systematik nur sehr bedingt derjenigen der geltenden Norm entspricht.

Für den didaktischen Kontext ist dieser Befund von enormer Wichtigkeit: Er verdeutlicht, dass Orthografierwerb nicht als schrittweise Übernahme eines Normenkatalogs modellierbar ist, sondern als Modifikation eines kognitiven Systems, aus dem das aktuelle graphematische Verhalten generiert wird. Ziel der didaktischen Intervention muss es daher sein, systematisch modifizierend auf dieses System einzuwirken, um eine zunehmende Passung mit der geltenden Norm zu erwirken. Die dafür erforderlichen methodischen Schritte sind aus dem jeweils aktuellen Ist-Zustand abzuleiten, für dessen Charakteristika die dargestellte Datenlage wesentliche, aber noch nicht erschöpfend untersuchte Anhaltspunkte liefert.

### Personeneffekte

Im Folgenden wird untersucht, inwiefern sich die Kompetenz der überwiegend aus der Sekundarstufe stammenden Probanden durch Persönlichkeitsvariablen wie Klassenstufe, Geschlecht und Schulart modellieren lässt. Erneut wurde dazu ein Regressionsmodell geschätzt, dessen Ergebnisse hier aber nur in Auszügen dargestellt werden, da das Modell lediglich als Pilotstudie einer deutlich detaillierteren Erhebung im Rahmen des FORTE-Projekts zu werten ist.

Tab. 5: Statistische Kennwerte des Personenparameters

	<b>N</b>	<b>Minimum</b>	<b>Maximum</b>	<b>Mittelwert</b>	<b>Std.abweichung</b>
PP	8985	-0,916180	6,000565	2,05538147	0,935793365

Eine Regressionsanalyse unter Einbezug der Variablen „Land“, „Bundesland/Kanton“, „Klassenstufe“, „Schulart“ und „Geschlecht“ zeigt ein höchst signifikantes, aber nur bedingt befriedigendes Ergebnis. Die Aufklärungsrate des Modells liegt bei lediglich 12,4% (angepasstes  $R^2 = ,124$ ), zeigt also einen erheblichen Anteil unaufgeklärter Varianz.<sup>13</sup> Offenbar werden wesentliche Persönlichkeitsfaktoren, die Auswirkungen auf die orthografische Kompetenz der Probanden haben, im Modell noch nicht erfasst. Da der Großteil der Einflussvariablen kategorial war, wurde eine programmgestützte automatische Datenvorbereitung vorgenommen, bei der die Kategorien für eine maximale Assoziation mit der abhängigen Variable zusammengeführt wurden.

<sup>13</sup> Um die Modellqualität zu erhöhen, wurden auch andere lineare Modelle geschätzt, so insbesondere generalisierte lineare Mischmodelle (GLMM), bei denen Land, Bundesland und Klassenzugehörigkeit als zufällige Effekte einbezogen wurden. Die Qualitätskriterien zeigten eine sehr schwache Verbesserung der Modellqualität an, die jedoch nicht hinreichte, um die Komplexitätszunahme des Modells zu rechtfertigen.

Der Prädiktoreinfluss der Variablen ist zusammen mit dem zugehörigen Signifikanzniveau und der Anzahl der automatisch erzeugten Kategorien in Tabelle 6 abgetragen.

Tab. 6: Koeffizienten der Regressionsanalyse (kumulierte Kategorien),  
abhängige Variable: Personenparameter

Prädiktor	Wichtigkeit	Signifikanz	Anzahl Kategorien
Schulart	.535	,000	7
Klassenstufe	.378	,000	- (Intervall)
Bundesland/Kanton	.053	,000	7
Geschlecht	.030	,000	2
Land	.005	,069	3

Aus der Analyse resultiert, dass die Schulart im Modell den bedeutendsten Prädiktor darstellt. Dazu wurden die 23 separierten Schulartenkategorien in sieben Einzelvariablen geclustert. Abbildung 2 gibt einen Überblick über die Kompetenzmittelwerte der Schularten und offenbart, dass die Kompetenzverteilung keineswegs durchgängig mit der Höhe des jeweils möglichen Schulabschlusses korreliert. Zwar weisen die meisten Schulen des sekundären und tertiären Bildungsbereichs erwartungsgemäß eine hohe mittlere Kompetenz auf, doch enthält das Diagramm auch erwartungswidrige „Ausreißer“, insbesondere beim sehr guten Abschneiden der Förderschule bzw. beim „Schlusslicht“ der Sekundar-, Ober- und Gesamtschulen mit gymnasialer Oberstufe. Denkbar ist, dass in diesen Schularten ein größerer Wert auf das Erlernen von Rechtschreibkompetenzen gelegt wird als in anderen. Angesichts der Zusammensetzung der Daten muss hier freilich auch mit Selektionseffekten gerechnet werden.

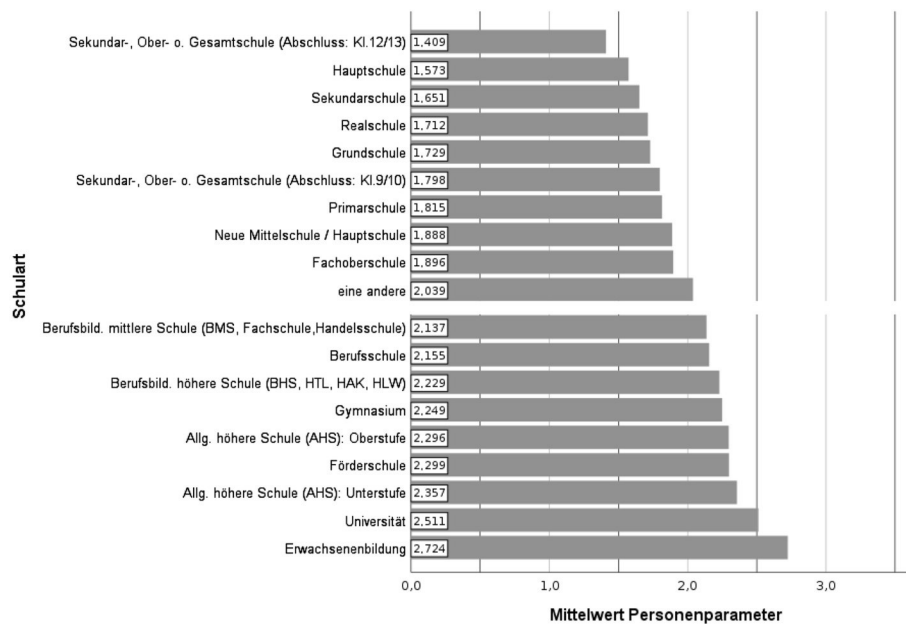


Abb. 2 Mittelwert des Personenparameters, verteilt auf die einzelnen Schularten ( $N \geq 10$ )

Den zweitgrößten Einfluss auf die Personenkompetenz zeigt mit einem Anteil von 37,8% die Variable Klassenstufe. In der Regressionsanalyse erreichte sie einen Koeffizienten von  $B = ,110$  (Std.fehler = ,006,  $p < ,001$ ). Er besagt, dass die Kompetenz der Getrennt- und Zusammenschreibung im Mittel pro Klassenstufe um 0,11 Kompetenzpunkte zunimmt, was bei einer Streuung von  $\sigma = 0,936$  bemerkenswert wenig ist. Die Befundlage deckt sich mit anderen statistischen Erhebungen (so analog etwa bei Müller 2007: 194, Müller 2016: 319).<sup>14</sup> Darüber hinaus zeigt Abbildung 3, dass der Kompetenzanstieg innerhalb der Klassenstufen nicht linear verläuft,<sup>15</sup> sondern nach Abschluss der Sekundarstufe einen deutlichen Abwärtsknick aufweist, der möglicherweise eine diskontinuierliche Kompetenzentwicklung anzeigt (vgl. dazu Müller 2016: 338–343). Auch Mesch (2015: 273) berichtet aus Untersuchungen des frühen GZS-Erwerbs von charakteristischen Diskontinuitäten in

<sup>14</sup> Ein Vergleich mit den großen Fehlerstatistiken im Vorfeld der Rechtschreibreform (etwa Riehme/Heidrich 1970, Menzel 1985, Zimmermann/Heckel 1986 u. a.) ist schwer, weil in diesen Untersuchungen lediglich das Verhältnis der Fehlerarten untereinander angegeben ist. Auch in jüngeren Arbeiten mit anspruchsvollerer statistischer Methodik (etwa Thomé/Gomolka 2007) spielt GZS oft eine untergeordnete Rolle.

<sup>15</sup> Sollten sich die Hinweise auf einen nicht-linearen Entwicklungsverlauf der GZS-Kompetenz weiter verdichten, stünde freilich auch die Suffizienz der linearen Regression an sich in Frage, worauf ein anonymer Gutachter zu Recht hinweist. Sie ist daher als ein erster Schritt der Modellbildung zu betrachten, der möglicherweise in weiteren Untersuchungsschritten durch komplexere Modelle abgelöst werden muss.

individuellen Erwerbsverläufen und kritisiert die Vorstellung einer linearen Kompetenzsteigerung (vgl. ebd.: 276).

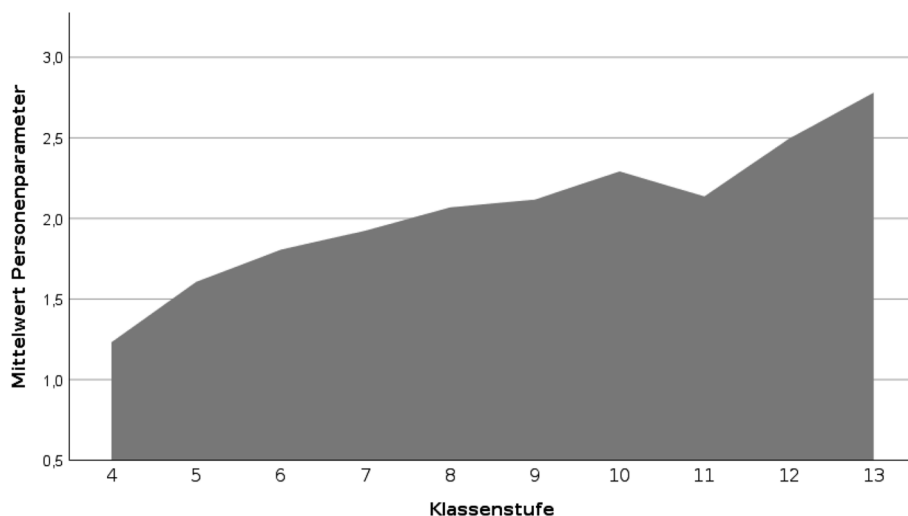


Abb. 3 Mittelwert des Personenparameters verteilt auf die Klassenstufen

Nur noch am Rande erwähnenswert sind die letzten drei Effekte auf die Personenkompetenz, die zusammen gerade einmal 9% zur Varianzaufklärung beitragen. 5,3% macht die regionale Herkunft nach Bundesland/Kanton aus, die damit immerhin noch statistisch signifikant wird, während der Effekt des Landes minimal und nicht mehr hinreichend abgesichert ist.

Die Variable Geschlecht weist einen Koeffizienten von  $B = -,107$  (Std.fehler: ,019,  $p < ,000$ ) auf und indiziert damit einen Leistungsvorsprung der Mädchen, der in etwa in der Größenordnung eines Schuljahres liegt.<sup>16</sup> Auch dieser Befund deckt sich mit der bekannten Datenlage zur orthografischen Kompetenz (vgl. etwa Kowalski/Voss 2009: 34, Grund/Haug/Naumann 2004: 65, Scheele 2006: 159).

Insgesamt ergibt die Untersuchung der Personenkompetenz eine mit den existierenden Daten hochgradig kompatible, wenngleich nicht triviale Befundlage. Der mit zunehmender Klassenstufe ansteigende Kompetenzparameter und der typische Leistungsvorteil der Mädchen sind wenig überraschend. Genauerer Analyse bedarf hingegen die Kompetenzverteilung auf die verschiedenen Schularten sowie die unterschiedlichen deutschsprachigen Regionen. Schließlich müsste präziser untersucht werden, ob es weitere Hinweise für eine diskontinuierliche Kompetenzentwicklung gibt. Allerdings hängt der Erfolg weiterführender Untersuchungen davon ab, inwie-

<sup>16</sup> Freilich ist diese Angabe angesichts der Indizien für einen diskontinuierlichen Entwicklungsverlauf mit Vorsicht zu genießen. Dennoch kann sie bis zum Leistungsknick in der 10. Klasse als vorläufiges Richtmaß für die Größenordnung des Geschlechterunterschiedes hinreichen.

fern es gelingt, den Anteil aufgeklärter Varianz des Kompetenzparameters durch Einbezug weiterer Persönlichkeitsvariablen wie z. B. der Herkunftssprache zu erhöhen. Die zu diesem Zweck angestrebte Datenerhebung im Rahmen des FORTE-Projekts ist derzeit in Arbeit.

## Schluss

Nach wie vor gibt die GZS nicht nur theoretische, sondern auch empirische Rätsel auf, die im Rahmen dieses Beitrages weniger gelöst als in ihrer Tragweite erfassbar gemacht worden sind. Die Verhaltensdaten der Versuchspopulation zeigen eine bemerkenswerte Systematik, die sich in starken Korrelationen mit sehr verschiedenen grammatischen Merkmalen niederschlägt. Anders als in anderen Teilbereichen der Orthografie fällt es angesichts der aktuellen Forschungslage dennoch schwer, sämtliche Phänomene in ein übergreifendes Erwerbsmodell zu integrieren. Dazu bedarf es neben einer genaueren Untersuchung der Einflüsse auf die orthografische Kompetenz insbesondere eines empirisch fundierten Kompetenzstufenmodells, mit dem sich die auf den Gedanken Bredels (2006) basierende Erwerbsmodellierung Meschs (2015) prüfen ließe.

Fest steht hingegen bereits jetzt, dass ein ausschließlich auf morphosyntaktischen Kriterien aufbauendes Modell, wie es Fuhrhop (2015a, b) und Jakobs (2005) zur Erfassung der Regularitäten der deutschen GZS ausarbeiten, nicht hinreicht, um neben der orthografischen Norm auch den schulischen Erwerbsverlauf adäquat zu modellieren. Grammatische Regularitäten, wie sie sich in typischen Betonungsmustern des Deutschen oder in semantischen und topologischen Proben manifestieren, mögen für die Beschreibung der Norm entbehrlich sein, werden von den Lernenden aber augenscheinlich umfassend genutzt, um das eigene graphematische Verhalten zu steuern. Fraglich bleibt schließlich, ob die Daten auch einer grundlegenden Überarbeitung der gültigen Norm das Wort reden.

## Literatur

- Amtliche Regelung (2006): In: <http://www.rechtschreibrat.com/> [mit den Nachträgen aus dem Bericht 2010]
- Augst, Gerhard/Dehn, Mechthild (2007): Rechtschreibung und Rechtschreibunterricht. Eine Einführung für Studierende und Lehrende aller Schulformen. 3., überarbeitete und aktualisierte Auflage. Stuttgart, Leipzig: Klett.
- Blatt, Inge (2015): Wie man Wortgrenzen entdeckt – GZS im Kernbereich und in der Peripherie. In: Bredel, Ursula/Reißig, Tilo (Hg.): Weiterführender Orthografieerwerb. 2., korrigierte Auflage. Baltmannsweiler: Schneider Verlag Hohengehren. S. 571–584.
- Bredel, Ursula (2006): Die Herausbildung des syntaktischen Prinzips in der Historiogenese und in der Ontogenese der Schrift. In: Bredel, Ursula/Günther, Hartmut (Hg.): Orthographietheorie und Rechtschreibunterricht. Tübingen: Niemeyer. S. 139–162.

- Eisenberg, Peter (2017): Deutsche Orthografie. Regelwerk und Kommentar. Verfasst im Auftrag der Deutschen Akademie für Sprache und Dichtung. Berlin, Boston: de Gruyter.
- Fuhrhop, Nanna (2015a): Orthografie. Vierte, aktualisierte Auflage. Heidelberg: Winter.
- Fuhrhop, Nanna (2015b): System der Getrennt- und Zusammenschreibung. In: Bredel, Ursula; Reißig, Tilo (Hrsg.): Weiterführender Orthografieerwerb. 2., korrigierte Auflage. Baltmannsweiler: Schneider Verlag Hohengehren. S. 107–128.
- Grund, Martin/Haug, Gerhard/Naumann, Carl Ludwig (2004): DRT 5. Diagnostischer Rechtschreibtest für 5. Klassen. Manual. 2. aktualisierte Aufl. in neuer Rechtschreibung, Göttingen: Hogrefe.
- Güthert, Kerstin (2011): Zur Neuregelung der deutschen Rechtschreibung ab 1. August 2006. Nachtrag Juli 2011 = Sprachreport. Extraausgabe Juli 2011.
- Heller, Klaus: (1996): Rechtschreibreform. Eine Zusammenfassung. = Sprachreport. Extraausgabe Juli 1996.
- Jacobs, Joachim (2005): Spatien. Zum System der Getrennt- und Zusammenschreibung im heutigen Deutsch. Berlin, New York: de Gruyter.
- Kowalski, Kerstin/Voss, Andreas (2009): Die IGLU-Ergänzungsstudie 2006 zur Rechtschreibkompetenz von Viertklässlern. Eine vergleichende Analyse von drei Tests. In: Valtin, Renate/Hofmann, Bernhard (Hg.): Kompetenzmodelle der Orthographie. Empirische Befunde und förderdiagnostische Möglichkeiten. Deutsche Gesellschaft für Lesen und Schreiben, Berlin: DGLS. S. 26–38.
- Menzel, Wolfgang (1985): Rechtschreibunterricht. Praxis und Theorie; aus Fehlern lernen. Seelze (Praxis Deutsch. Beiheft 69).
- Mesch, Birgit (2015): Konzepte des Erwerbs der Getrennt- und Zusammenschreibung. In: Bredel, Ursula/Reißig, Tilo (Hg.): Weiterführender Orthografieerwerb. 2., korrigierte Auflage. Baltmannsweiler: Schneider Verlag Hohengehren. S. 268–295.
- Müller, Hans-Georg (2007): Zum „Komma nach Gefühl“. Implizite und explizite Kommakompetenz von Berliner Schülerinnen und Schülern im Vergleich. Frankfurt am Main [u. a.]: P. Lang.
- Müller, Hans-Georg (2016): Der Majuskelgebrauch im Deutschen. Groß- und Kleinschreibung theoretisch, empirisch, ontogenetisch. Berlin: de Gruyter.
- Nottbusch, Guido/Jonischkait, Julia (2015): Einzeluntersuchungen zur GKS, GZS und Interpunktion. In: Bredel, Ursula/Reißig, Tilo (Hg.): Weiterführender Orthografieerwerb. 2., korrigierte Auflage. Baltmannsweiler: Schneider Verlag Hohengehren. S. 164–187.
- Riehme, Joachim/Heidrich, Marianne. (1970): Die Fehlerursachen analysieren, um die Schwerpunkte der Übungen zu erkennen. Über die Verteilung von orthographischen und grammatischen Fehlern in den Klassen 4 bis 10 und die methodischen Schlussfolgerungen. In: Deutschunterricht 6 und 7/8. S. 346–355 und 426–433.
- Rost, Jürgen (2004): Lehrbuch Testtheorie – Testkonstruktion. 2. überarb. Aufl. Bern [u. a.]: Huber Universitäts- und Landesbibliothek.
- Scheele, Veronika (2006): Entwicklung fortgeschrittener Rechtschreibfertigkeiten. Ein Beitrag zum Erwerb der „orthographischen“ Strategie, Frankfurt am Main: Lang.
- Thomé, Günther/Gomolka, Jens (2007): Rechtschreiben. In: Beck, Bärbel/Klieme, Eckhard (Hg.): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International): Weinheim: Beltz. S. 140–146.
- Valtin, Renate et al. (2003): Orthographische Kompetenzen von Schülerinnen und Schülern der vierten Klasse. In: Bos, Wilfried et al. (Hg.): Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich, Münster: Waxmann. S. 227–264.



Zimmermann, Friedrich/Heckel, Brigitte (1986): Untersuchungen zu Normverstößen bei der Groß- und Kleinschreibung. In: fi Forschungsinformation. Analysen orthographischer Leistungen von Schülern der POS der DDR, Pädagogische Hochschule Ernst Schneller Zwickau. S. 21–35.

#### **Anschrift der Verfasser**

*PD Dr. Hans-Georg Müller, Universität Potsdam, Am Neuen Palais 10, 14469 Potsdam, hgmuelle@uni-potsdam.de*

*Prof. Dr. Matthis Kepser, Universität Bremen, Bibliothekstraße 1, 28357 Bremen, kepsen@uni-bremen.de*

*Dr. Stefan Schallenberger, Universität Bremen, Bibliothekstraße 1, 28357 Bremen, schallenberger@uni-bremen.de*